AMERICAN JOURNAL OF PHYSICAL ANTHROPOLOGY 111:25–44 (2000)

Identifying Conservation Units Within Captive Chimpanzee Populations

AMOS S. DEINARD,*[1] AND KENNETH KIDD[2]
[1]*Department of Anthropology, Yale University, New Haven, Connecticut 06520*
[2]*Department of Genetics, Yale University, New Haven, Connecticut 06520*

KEY WORDS chimpanzees; conservation units; captive management; genetic diversity

ABSTRACT One of the primary objectives in the captive management of any endangered primate is to preserve as much as possible the genetic diversity that has evolved and still exists in wild gene pools. The rationale for this is based on the theoretical understanding of the relationship between genetic diversity and fitness in response to selection. There remains little consensus, however, as to the type of genetic data that should be used to monitor captive populations. In order to develop a deeper understanding of the degree and nature of genetic diversity among "wild" chimpanzee gene pools, as well as to determine if one type of genetic data is more useful than others, DNA sequence data were generated at three unlinked, nonrepetitive nuclear loci, one polymorphic microsatellite, and the mitochondrial D-loop for 59 unrelated common and pygmy chimpanzees. The results suggest that: 1) data from nuclear loci can be used to differentiate common chimpanzee subspecies; 2) pygmy chimpanzees may have less genetic diversity than common chimpanzees; 3) shared microsatellite alleles do not always indicate identity by descent; and 4) nonrepetitive loci provide unique insights into evolutionary relationships and provide useful information for captive management programs. Am J Phys Anthropol 111:25–44, 2000.

One of the primary objectives in the captive management of endangered primates is to preserve as much as possible the genetic diversity that has evolved and still exists in wild gene pools (Chesser et al., 1980; Schonewald-Cox et al., 1983). The rationale for this is based on the theoretical understanding of the relationship between genetic diversity and fitness in response to selection: loss of genetic diversity increases autozygosity and reduces a population's ability to "adapt" to future selective pressures (Fisher, 1930; Chesser et al., 1980; Lacy, 1987; Mitton, 1997). Although this relationship is well understood, and is the basis for most captive breeding programs, there remains little consensus as to the type of genetic data that should be used to monitor captive populations. Ideally, of course, management decisions should be based on the most comprehensive data available, including biogeographical, ecological, behavioral, morphological, and genetic information. Given the limited resources for generating molecular data, however, it becomes necessary to examine what types of data are the most appropriate in that they will generate maximum information on the nature and degree of genetic diversity within a captive population.

Grant Sponsor: Leakey Foundation; Grant Sponsor: Wenner-Gren Foundation for Anthropological Research; Grant Sponsor: National Science Foundation; Grant number: SBR 9315871.

A previous version of this paper was awarded the 1995 Sherwood Washburn Prize for a student-presented paper to A.S.D.

*Correspondence to: Amos S. Deinard, Dean's Office-Student Programs, School of Veterinary Medicine, University of California, Davis, CA 95616. E-mail: asdeinard@ucdavis.edu

Received 20 May 1998; accepted 17 August 1999.

In order to develop a deeper understanding of these issues, our laboratory has generated DNA sequence data at: 1) three unlinked, noncoding, single-copy, nuclear loci (a ~1.0-kb intergenic sequence located approximately 2.5 kb upstream from the Homeobox B6 (HOXB6) transcription start site at (human chromosome) 17q21–q22; a ~1.1-kb intron within the Apolipoprotein B gene (APOB, intron "AB" between exons 28–29) located at (human chromosome) 2p24–p23; and a ~750-bp region of the X chromosome pseudoautosomal boundary locus (PABX) located at Xp22 and a ~940-bp region of the Y chromosome pseudoautosomal boundary locus (PABY) located at Yp11); 2) one highly polymorphic short tandem repeat (STR or microsatellite) at HOXB6; and 3) approximately 340 bp of the hypervariable region (HVI) of the mitochondrial control region (D-loop) for 18 "wild-caught" pygmy chimpanzees or bonobos (*Pan paniscus*) and 41 unrelated "wild-caught" common chimpanzees.

The results from our comparisons not only confirm previous studies that have detected a high level of intraspecific variation among chimpanzees (Ruano et al., 1992; Ely et al., 1992; Morin et al., 1994; Ruvolo et al., 1994; Rubinsztein et al., 1995), but they suggest that the analysis of single-copy, unlinked, noncoding, nonrepetitive nuclear loci for nucleotide variation (e.g., point mutations, small insertions/deletions) may yield specific insights into the evolutionary relationships among individuals that analyses of mitochondrial DNA or STR loci cannot provide. This argument is based on our belief that the analysis of such nuclear loci may provide a greater (and underutilized) resource from which to identify "conservation management units" (Moritz, 1994a, 1995).

MATERIALS AND METHODS

Detection of DNA sequence variation

The HOXB6 and APOB loci were first amplified by polymerase chain reaction (PCR) (Saiki et al., 1985) and checked for product quantity and size variation by agarose gel electrophoresis. Denaturing gradient gel electrophoresis (DGGE) was then employed to detect nucleotide variation among samples (Fischer and Lerman, 1983; Lerman et al., 1984; Myers and Maniatis, 1986; Myers et al., 1987, 1989; PCR and DGGE conditions are presented in Appendix A). Detection via DGGE results from a gradient of denaturants within the gel which simulate a gradient of increasing temperature (Lerman et al., 1984). Since molecules with different sequences focus at different regions within the gel, DGGE of PCR amplified products is a highly sensitive method by which one can detect variations in DNA sequence differing by as little as a single base (Fischer and Lerman, 1983; Myers et al., 1987, 1989; Kidd and Ruano, 1992; Deinard and Kidd, 1998).

The next step in the analysis of the HOXB6 and APOB loci was to isolate and sequence all detected allelic variants. For individuals that were homozygous for a variant allele, that individual's amplified DNA was purified and used as the template for the sequencing reaction. This was always possible for the most common alleles. If a variant allele was only observed in individuals heterozygous for the variant (e.g., heterozygotes for the common allele and a less common variant), however, the homoduplex band for the variant allele was isolated by excising the band from the denaturing gradient gel and the DNA was purified (Sambrook et al., 1989). The isolated and purified DNA was reamplified and DGGE was performed on the amplified DNA. If DGGE demonstrated a single (variant) band, indicating a successful isolation of the variant allele, the remaining PCR product was purified and served as the template for the sequencing of the allelic variant.

DNA sequencing

Data from the pseudoautosomal X (PABX) and Y (PABY) boundaries as well as the mitochondrial D-loop were generated strictly from the direct sequencing of PCR-amplified products (see Appendix A). Although all individuals were sequenced for their mitochondrial D-loop, only male individuals were analyzed for the PABX and PABY loci. This approach prevented ambiguity within the DNA sequences generated: the mitochondrial genome is haploid, and males only have single copies of PABX and PABY.

For all loci, DNA sequencing was completed in both directions for any given DNA fragment examined. All HOXB6 and APOB PCR fragments analyzed via DGGE also served as templates for sequencing purposes, with all PCR primers doubling as sequencing primers. Consequently, the only sequencing-specific primers used were to generate sequence data at PABX and PABY. DNA sequencing was primarily accomplished by cycle sequencing with fluor-labeled terminators, using the ABI Prism Dye Terminator Cycle Sequencing Core Kit with AmpliTaq DNA polymerase, FS and an ABI 373 DNA-sequencer (Applied Biosystems, Inc.). If the automated DNA sequencing resulted in ambiguous sequences or if an allelic variant detected via DGGE could not be isolated by excising (due to identical DGGE migration of the homoduplex bands), the PCR product was manually sequenced (Sequenase PCR sequence kit, USB). In this manner, all sequence variants were unambiguously determined.

If two or more individuals could not be distinguished via DGGE heteroduplex analysis (Ruano and Kidd, 1992), it was assumed that they shared identical sequences, and thus a DNA sequence was not generated for each individual. This approach drastically cut down on the time and cost of analyzing the large number of individuals sampled for this project and demonstrates the usefulness of DGGE for large-scale sequence screening (Deinard and Kidd, 1998). For the mitochondrial D-loop, PABX, and PABY loci, all individuals analyzed were completely sequenced for the entire region of DNA examined.

Analysis of HOXB6 STR

The short tandem repeat (STR) located within the HOXB6 intergenic region analyzed for this project has been previously observed to be polymorphic in humans and great apes (Deinard et al., 1992; Deinard and Kidd, 1995). This STR, a dinucleotide (CA) repeat, was subsequently analyzed in all individuals independently from the DGGE analysis of the upstream and downstream HOXB6 flanking sequences (Fig. 1a), following the protocol and conditions previously published (Deinard et al., 1992; Deinard and Kidd, 1995).

Statistical analyses and methods

The last step of our analysis was to construct molecular haplotypes for the entire region of the sequence examined at HOXB6 and APOB. This was readily accomplished because each region was subdivided into overlapping fragments for the purposes of screening the region via DGGE (Fig. 1a,b). As a result, it was often obvious how any given detected nucleotide substitution was related to the sequence upstream and/or downstream from it. In fact, the only region that proved consistently problematic was HOXB6 STR. In some instances, it was obvious that specific repeat lengths were associated with a given nucleotide substitution up- or downstream. In those cases that remained ambiguous, however, PCR products of the entire (~1.0 kb) region were cloned (TA Cloning Kit, Invitrogen) and the cloned products were manually sequenced (Sequenase PCR sequence kit, USB). This allowed the precise DNA sequence to be determined, including the exact CA repeat length for any given HOXB6 allele. Since the mitochondrial genome is haploid, and only single copies of PABX and PABY were examined, no ambiguity existed in the haplotypes of these loci.

Once all the individual haplotype sequences for a specific locus were determined, they were aligned. This was accomplished using the Clustal algorithm of DNASTAR's MegAlign software (gap penalty, 10; gap length penalty, 10). From this alignment and with human and/or gorilla consensus sequences serving as outgroup(s), DNA sequence phylogenies (gene trees) were constructed using maximum parsimony (PAUP 3.0, using a heuristic simple stepwise addition; Swofford, 1990). Comparisons of the resulting allelic phylogenies allowed for inferences to be made concerning the evolutionary relationships between the chimpanzees sampled. In addition, several sequence statistics were calculated using DnaSP 2.52 (Rozas and Rozas, 1997) to compare the levels of genetic diversity (i.e., nucleotide diversity, π) within and between the chimpanzee populations and to test the mutational neutrality (i.e., Tajima's D and (Tajima, 1989) the D and F statistics of Fu and Li (1993), of each locus analyzed.

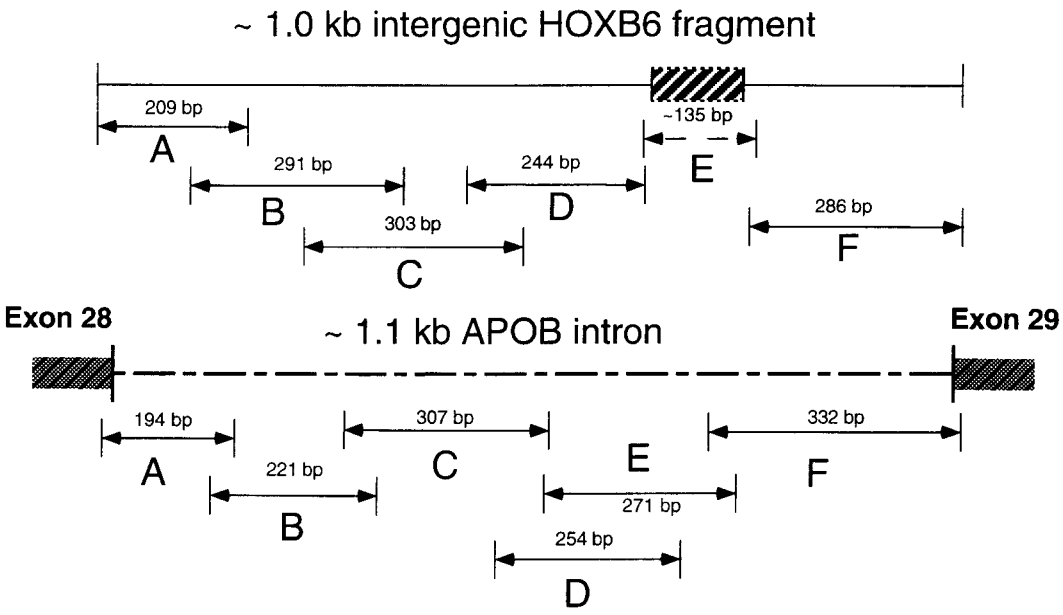

Fig. 1. **a:** Analytic strategy for the HOXB6 fragment. Five fragments, represented by the solid lines A–D and F, were individually analyzed via DGGE in order to screen the ~1.0-kb intergenic HOXB6 region for sequence variation. The size of each fragment (in bp) is indicated. Briefly, all PCR reactions were performed with 200 ng genomic DNA, 0.1 mM each primer, 0.5 U Taq polymerase, 50 mM each nucleotide in 100 µl reaction buffer (50 mM KCl, 1.5–3.75 mM $MgCl_2$, and 10 mM Tris, pH 8.4) for 30–35 cycles. The thermal profiles for each fragment analyzed, the $MgCl_2$ concentration of the PCR reaction buffer used, and the electrophoretic conditions are presented in Appendix A. Fragment E, represented by the broken line and hatched box, corresponds to the CA repeat and was independently analyzed. Primer pairs (forward; reverse) used for each fragment were as follows: fragment A, 5′-aaccaaggtgttgccagag-3′; 5′-@-tcagggcctcctatagagca-3′; fragment B, 5′-@-cctcacctcctttcggcg-3′; 5′-ggggagtgagcagtagtgctga-3′; fragment F, 5′-@-gcaaatatccctcctgaggtt-3′; 5′-atacacagcgggctcagc-3′. "@-" represents the "GC clamp" (the sequence 5′-CGCCCGCCGCGCCCCGCGCCCGGCCCGCCGCCCCCGCCC-3′), synthesized as part of the PCR primer. Primer pairs for fragments C, D, and E have been previously described (Ruano et al., 1992; Deinard et al., 1992; Deinard and Kidd, 1995). **b:** Analytic strategy for APOB intron. Six fragments, represented by the solid lines A–F, were individually analyzed in order to screen the ~1.1-kb APOB intron for sequence variation. The size of each fragment (in bp) is indicated. The basic PCR conditions for each APOB fragment were identical to those of the HOXB6 fragments; the thermal profiles for each fragment analyzed, the $MgCl_2$ concentration of the PCR reaction buffer used, and the electrophoretic conditions are presented in Appendix A. APOB primer pairs (forward; reverse) were as follows: fragment A, 5′-@-gggcatggatatggatgaag-3′; 5′-agaagaagttgttgcttaccgc-3′; fragment B, 5′-ttcagagattcccaaactagg-3′; 5′-@-ctgcctatggctgagcttg-3′; fragment C, 5′-taattctgggcaaatccagg-3′; 5′-@-cccagaatcccaacatcagt-3′; fragment D, 5′-@-catgtcatttacaagctctgcc-3′; 5′-gaaaactggaagaggagtggg-3′; fragment E, 5′-@-gttgggattctgggtctc-3′; 5′-tcatgtgtttgtttttatggttaa-3′; fragment F, 5′-taaatcctttctgggggcatg-3′; 5′-@-agaagctgcctcttcttccc-3′. "@-" represents the "GC clamp" (the sequence 5′-CGCCCGCCGCGCCCCGCGCCCGGCCCGCCGCCCCCGCCC-3′), synthesized as part of the PCR primer.

Individuals sampled

Eighteen bonobos (*Pan paniscus*) and 41 common chimpanzees (29 *Pan troglodytes verus*, 9 *Pan troglodytes troglodytes*, and 3 *Pan troglodytes schweinfurthii*) were sampled for this project. Seven of the *P. t. troglodytes* individuals originate from Gabon, one originates from Cameroon, and one originates from Equatorial Guinea, while 12 of the 29 *P. t. verus* individuals originate from Liberia (see Appendix B). All individuals were either "wild-caught" or were the offspring of wild-caught individuals unavailable for sampling (Schobert, 1991; Reinartz, 1994). In addition, all individuals were unrelated except for two bonobos, Maiko and Congo; the sire and dam of Maiko and the dam of Congo are full siblings (Reinartz, 1994). Consequently, Maiko and Congo may share a chromosome, while Maiko may have inherited the same chromosome twice. Accordingly, the number of unrelated chromosomes examined for each species was as follows: *P. troglodytes*, N = 82; *P. paniscus*, N = 34–36. The human and gorilla sequences that served as the outgroups for the

parsimony analyses represented consensus sequences generated from a larger data set (Deinard, 1997); the topologies of the chimpanzee gene trees using consensus human and gorilla sequences as outgroups did not vary from those generated using several human and gorilla sequences (Deinard, 1997).

RESULTS

HOXB6

The first locus examined for this project was the ~1.0-kb intergenic HOXB6 sequence. The evolutionary relationships among all observed HOXB6 haplotypes within *Pan*, excluding the STR, are depicted in Figure 2. The frequencies of each observed haplotype are presented in Table 1a. Several observations can be made from these data. The first is that except for one common haplotype (the haplotype "Trog 2") that is carried by all three common chimpanzee subspecies, all other HOXB6 haplotypes are unique to a single subspecies (Table 1a). A second observation is that bonobos carry fewer haplotypes than do common chimpanzees (5 vs. 8 haplotypes), although we note that *P. paniscus* demonstrate an equal degree of nucleotide diversity as do *P. troglodytes* (Table 2).

A third observation is that among the bonobos sampled, two unique clusters of haplotypes (consisting of haplotypes A, D, and E and haplotypes B and C) were detected (Fig. 2). Although by all definitions *P. paniscus* is a "good" species, supported by behavioral, genetic, and morphological data (Susman, 1984; Kano, 1986), from an evolutionary standpoint these haplotypes are no "closer" to each other than they are to certain common chimpanzee haplotypes: bonobo haplotype A and haplotype B (with frequencies of 64% and 25%, respectively; Table 1a) are separated by four mutational events, whereas both of these haplotypes are only two mutational events away from the common chimpanzee haplotype "Trog 2" (Fig. 2).

A fourth observation may be made concerning the STR (i.e., the CA repeat) that lies within the HOXB6 region analyzed. In contrast to the flanking sequence either immediately upstream or downstream (fragments A–D, F; Fig. 1a), the HOXB6 STR (fragment E; Fig. 1a) proved to be highly polymorphic, with STR alleles ranging in size from 126–150 bp (Table 1b). Not only were more alleles detected at the STR when compared to the surrounding (nonrepetitive) sequence (Table 1b), but the STR has a higher degree of observed heterozygosity when compared to the surrounding sequence (Table 1c). These data demonstrate that the surrounding flanking sequence has less intraspecific genetic variation than the STR and imply

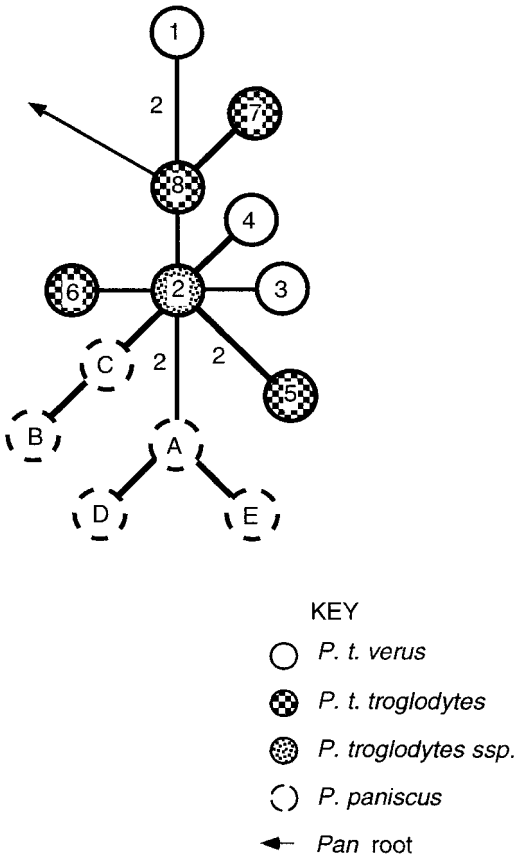

Fig. 2. The maximum parsimony network (representing 1 of 5 equally parsimonious trees) for the observed haplotypes (excluding the STR) at HOXB6. Each circle represents an observed haplotype: *P. troglodytes* haplotypes are represented by closed circles and numbered, while *P. paniscus* haplotypes are represented by broken circles and lettered. Mutational distances between haplotypes greater than one are indicated. The *Pan* root is indicated by the arrow, and represents the branch leading to the outgroup (i.e., *Gorilla* or *Homo*). The observed frequencies for each *Pan* haplotype, as defined by the assigned numbers and letters, are presented in Table 1a. The DNA sequences for each observed haplotype, excluding variation at the STR, have been deposited in Genbank (accession nos. AF116786–AF116798).

TABLE 1a. Frequencies of observed HOXB6 haplotypes[1]

P. paniscus, observed haplotypes	Number of chromosomes with observed haplotype	Frequency
Paniscus A	N = 23	64.0%
Paniscus B	N = 9	25.0%
Paniscus C	N = 1	2.8%
Paniscus D	N = 2	5.6%
Paniscus E	N = 1	2.8%

	Number of chromosomes with observed haplotype			
P. troglodytes, observed haplotypes	*verus*	*troglodytes*	*schweinfurthii*	Frequency
Trog 1	46			55.1%
Trog 2	10	8	6	29.3%
Trog 3	1			1.2%
Trog 4	1			1.2%
Trog 5		2		2.4%
Trog 6		6		7.3%
Trog 7		1		1.2%
Trog 8		1		1.2%

[1] Each haplotype corresponds only to the nonrepetitive HOXB6 sequence and does not include the HOXB6 STR. The evolutionary relationships among HOXB6 haplotypes are presented in Figure 2.

TABLE 1b. Observed haplotypes with their associated STR alleles

Observed haplotypes	STR alleles observed with each HOXB6 haplotype[1]												
	126	128	130	132	134	136	138	140	142	144	146	148	150
P. paniscus													
Paniscus A				x			x		x	x	x		
Paniscus B	x												
Paniscus C		x											
Paniscus D							x	x					
Paniscus E							x						
P. troglodytes													
Trog 1	x	x	x	x	x	x	x			x			
Trog 2			x	x	x	x		x	x	x	x		
Trog 3			x										x
Trog 4							x						
Trog 5							x						
Trog 6							x	x					
Trog 7									x				
Trog 8										x			

[1] As defined by the upstream and downstream sequence flanking the STR. See Materials and Methods.

TABLE 1c. Number of alleles and observed heterozygosity between the HOXB6 STR and its flanking sequences

	Number of alleles		Observed heterozygosity	
Species	CA repeat	Flanking	CA repeat	Flanking
P. troglodytes	12	8	67%	44%
P. paniscus	8	5	89%	50%

that each haplotype (as defined by the up- and downstream flanking sequence) has several CA repeat alleles associated with it (Table 1b).

APOB, pseudoautosomal boundaries, and the mitochondrial D-loop

The second locus examined for the project was the ~1.1 kb Apolipoprotein B intron. The evolutionary relationships among all observed APOB haplotypes within *Pan* are depicted in Figure 3. Although common chimpanzees showed a high degree of intraspecific variation at this locus (with 17 identified haplotypes), the overall pattern of the evolutionary relationships was similar to the pattern observed at HOXB6: common chimpanzee subspecies were observed to have unique haplotypes, except for four of the *P. t. schweinfurthii* haplotypes that clustered with the *P. t. troglodytes* haplotype "Trog 3" (Fig. 3, Table 3). In addition, at the APOB locus bonobos had fewer haplotypes (3 vs. 17) and a 10-fold difference in nucleotide diversity as compared to common chimpanzees (Tables 2 and 3).

TABLE 2. Sequence statistics for each locus[1]

Locus and statistic	Sample				
	Pt	Pp	Ptt	Ptv	Pts
HOXB6 (977 bp)					
No. of sequences examined	82	36	18	58	6
No. of polymorphic sites	8	6	5	5	0
π (in percent)	0.184	0.183	0.124	0.11	0
Tajima's *D*	na	0.643*	−0.532*	−0.020*	na
Fu and Li's D test	na	.382*	0.389*	−0.936*	na
F test	na	0.548*	0.152*	−0.760*	na
APOB (1,060 bp)					
No. of sequences examined	82	36	18	58	6
No. of polymorphic sites	17	2	15	2	5
π (in percent)	0.152	0.015	0.285	0.033	0.157
Tajima's *D*	na	−1.28*	−1.16*	−0.338*	−1.34*
Fu and Li's D test	na	−0.835*	−2.20**	−0.835*	−1.29*
F test	na	−1.12*	−2.27**	−1.12*	−1.53*
PABX (719 bp)					
No. of sequences examined	20	11	1	17	2
No. of polymorphic sites	6	1	0	4	2
π (in percent)	0.202	0.025	na	0.162	0.278
Tajima's *D*	na	−1.13*	na	−0.055*	na
Fu and Li's D test	na	−1.39*	na	1.16*	na
F test	na	−1.54*	na	0.974*	na
PABY (882 bp)					
No. of sequences examined	19	11	1	16	2
No. of polymorphic sites	5	2	0	0	0
π (in percent)	0.111	0.100	na	na	0
Tajima's *D*	na	0.85*	na	na	na
Fu and Li's D test	na	0.966*	na	na	na
F test	na	1.10*	na	na	na
mt *D*-loop (342 bp)					
No. of sequences examined	41	17	9	29	3
No. of polymorphic sites	108	53	45	74	13
π (in percent)	7.75	3.87	5.59	5.30	2.56
Tajima's *D*	na	−0.671*	0.470*	−0.455*	na
Fu and Li's D test	na	0.246*	0.795*	−0.238*	na
F test	na	−0.067*	0.870*	−0.376*	na

[1] Statistics are presented for each sample group compared: *P. troglodytes* ssp. (Pt), *P. paniscus* (Pp), *P. t. troglodytes* (Ptt), *P. t. verus* (Ptv), and *P. t. schweinfurthii* (Pts). Nucleotide diversity (π) corresponds to the average number of nucleotide differences between two sequences (Nei, 1987, eq. 10.5 or 10.6) and is presented as a percentage. Mutational neutrality for each locus was calculated using Tajima's *D* statistic and Fu and Li's *D* and *F* statistics (with outgroup) (Tajima, 1989; Fu and Li, 1993). The statistic was not calculated (na) for *P. troglodytes* ssp. (Pt), since such a group does not represent a true breeding population given the geographic isolation of each subspecies, or in those cases where there were too few haplotypes or no detected polymorphisms. A *D*-loop sequence was not generated for the bonobo Lady.
* $P > 0.10$.
** $0.10 > P > 0.05$.

Similar evolutionary relationships were observed from the analysis of the pseudoautosomal boundary sequence data, both on the X chromosome (PABX) and the Y chromosome (PABY). Although we note that only a single *P. t. troglodytes* and two *P. t. schweinfurthii* individuals were compared at these loci, the data nevertheless suggest the existence of subspecific clustering among common chimpanzees (Fig. 4, Table 4). A similar pattern of chimpanzee subspecies clustering is observed if all the nuclear sequence data are combined in a "total (nuclear) evidence approach" (Huelsenbeck et al., 1996) and used in the phylogenetic analyses (Fig. 5).

DISCUSSION

Genetic data demonstrating the clustering of common chimpanzee subspecies is not new. Morin et al. (1994) were able to differentiate subspecies based on the overall similarity (genetic distance) of mitochondrial D-loop sequences. Comparing the individual chimpanzees sampled for this project for the identical region of mitochondrial DNA (i.e., D-loop) reveals an identical pattern of subspecific clustering (Fig. 6). Given that these analyses of mitochondrial DNA (Morin et al., 1994; this project) reveal a similar pattern of intraspecific variation as the analyses of the nuclear loci examined for this project, one could conclude that one element

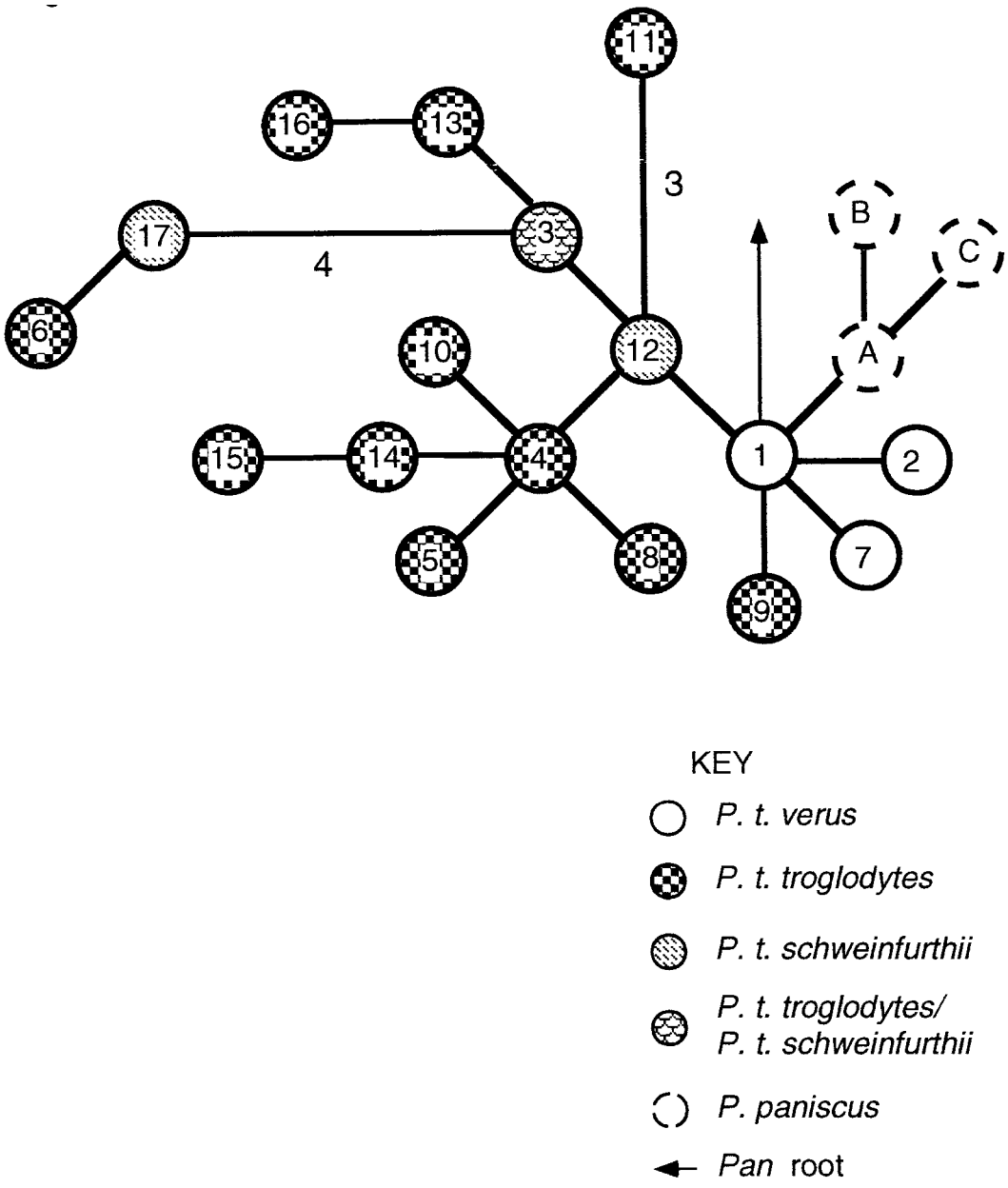

Fig. 3. The maximum parsimony network (representing 1 of 84 equally parsimonious trees) for the observed haplotypes at APOB. Each circle represents an observed haplotype: *P. troglodytes* haplotypes are represented by closed circles and numbered, while *P. paniscus* haplotypes are represented by broken circles and lettered. Mutational distances between haplotypes greater than one are indicated. The *Pan* root is indicated by the arrow and represents the branch leading to the outgroup (i.e., *Gorilla* or *Homo*). The observed frequencies for each *Pan* haplotype, as defined by the assigned numbers and letters, are presented in Table 3.

of chimpanzee captive management and conservation genetics (i.e., the subspecific classifying of captive individuals) should focus on the analysis of mitochondrial DNA. Such a conclusion would not be without merit: the mitochondrion is a useful locus from which to quickly and efficiently produce molecular data because it is a small (~16 kb), ubiqui-

TABLE 3. Frequencies of observed APOB haplotypes[1]

P. paniscus, observed haplotypes	Number of chromosomes with observed haplotype	Frequency
Paniscus A	N = 33	91.7%
Paniscus B	N = 2	5.5%
Paniscus C	N = 1	2.8%

	Number of chromosomes with observed haplotype			
P. troglodytes, observed haplotypes	*verus*	*troglodytes*	*schweinfurthii*	Frequency
Trog 1	46			53.6%
Trog 2	11			13.4%
Trog 3		5	4	11%
Trog 4		2		2.4%
Trog 5		2		2.4%
Trog 6		1		1.2%
Trog 7	1			1.2%
Trog 8		1		1.2%
Trog 9		1		1.2%
Trog 10		1		1.2%
Trog 11		1		1.2%
Trog 12			1	1.2%
Trog 13		1		1.2%
Trog 14		1		1.2%
Trog 15		1		1.2%
Trog 16		1		1.2%
Trog 17			1	1.2%

[1] The evolutionary relationships among APOB haplotypes are presented in Figure 3.

tous, maternally inherited unit that undergoes no recombination, and the complete chimpanzee mitochondrial sequence is known (Horai et al., 1995; Arnason et al., 1996), thus allowing for PCR and sequencing primers to be easily synthesized. The analysis of mitochondrial DNA may also yield a greater number of informative characters than the analysis of a nuclear locus (Moritz, 1994b; Avise, 1995), because mitochondrial DNA "evolves" faster than nuclear DNA (Brown et al., 1982).

Some of the same reasons that make the mitochondrion an attractive locus, however, also have the potential to limit the conclusions that can be drawn from its analysis. Primarily, comparisons based on mtDNA may not be representative of the true genetic relations among populations, as we think of them based on nuclear DNA. Although mitochondrial analyses may allow for individual chimpanzees to be subspecifically typed (e.g., Morin et al., 1992), they may nevertheless provide little information concerning the species-specific, subspecies-specific, or even population-specific conclusions that are the primary focus of most captive management or conservation programs. Even though few, if any, conservation programs purport to exclusively rely on mitochondrial data to make their (genetic) management decisions, it is worth noting that the majority of genetic studies on captive or wild primates, including chimpanzees, often limit their analyses to the mitochondrial D-loop. Although the high mitochondrial copy number partially explains why it remains the best locus to examine for generating data on (wild) individuals that cannot readily be invasively sampled (Gagneux et al., 1997), it does not entirely explain the abundance of primate studies that compare only mitochondrial genes despite having access to adequate (i.e., high molecular weight) DNA that could be used for comparing independent nuclear genes (e.g., Ruvolo et al., 1994; Hoelzer et al., 1994; Zhi et al., 1996; James et al., 1997; Tagliaro et al., 1997). Consequently, it may be worthwhile to briefly review several reasons why comparisons based solely on mtDNA may not be representative of the true genetic relations among populations, as we think of them based on nuclear DNA.

The first reason is that a population may maintain an equal degree of nuclear variation when compared to other populations within the species' range, even though it may have a limited number of mitochondrial DNA haplotypes. Consequently, mitochondrial data need not accurately duplicate the pattern of variation as observed from nuclear

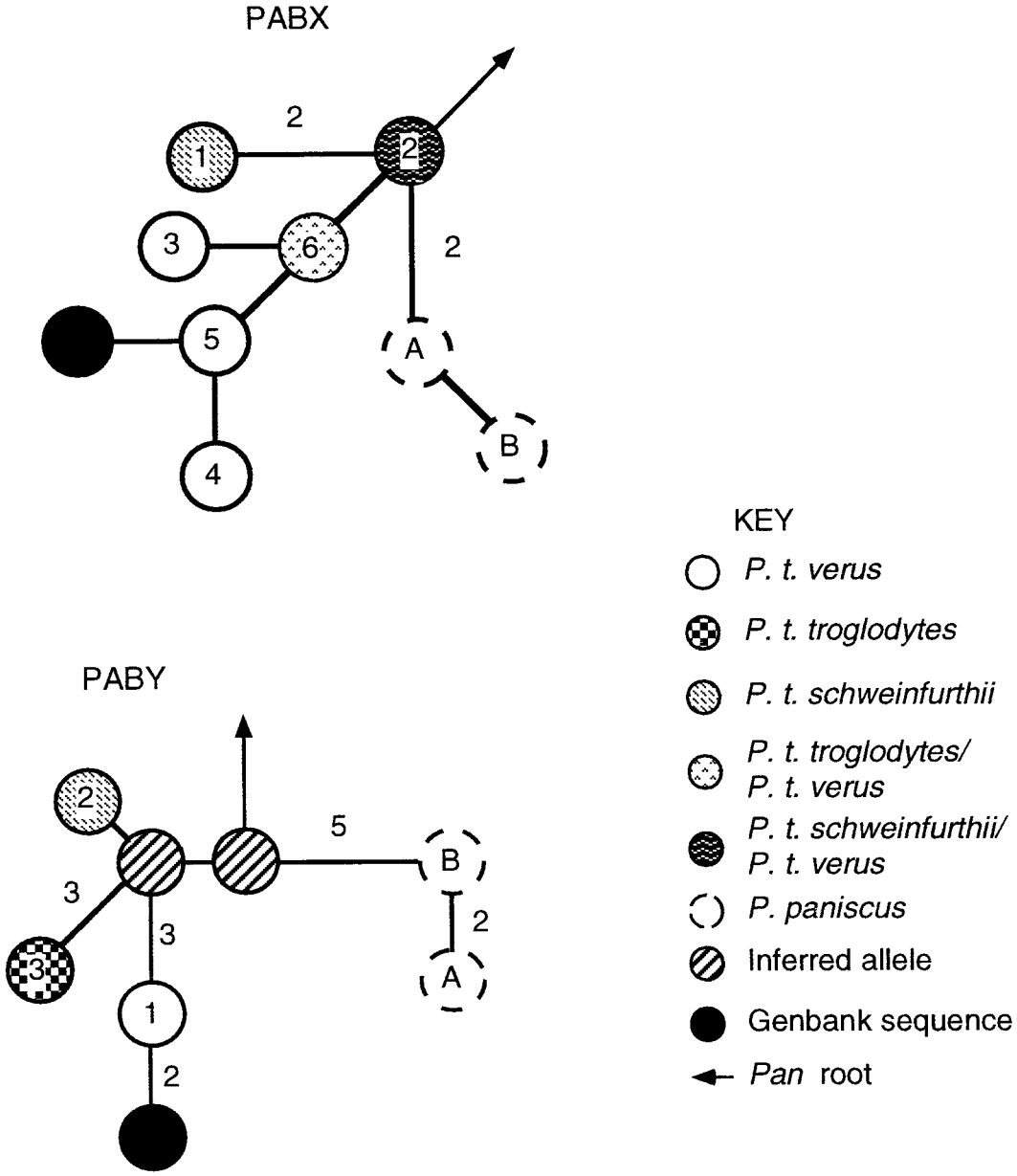

Fig. 4. The maximum parsimony networks for the haplotypes observed at PABX and PABY. Each circle represents an observed haplotype: *P. troglodytes* haplotypes are represented by closed circles and numbered, while *P. paniscus* haplotypes are represented by broken circles and lettered. Solid circles represent the common chimpanzee haplotypes defined by the PABX and PABY Genbank sequences (X53804 and X53803, respectively) that were used for comparative purposes, while hatched circles represent inferred ancestral haplotypes. Mutational distances greater than one are indicated. The *Pan* root is indicated by the arrow and represents the branch leading to the outgroup (i.e., *Gorilla* or *Homo*). The observed frequencies for each *Pan* haplotype, as defined by the assigned numbers and letters, are presented in Table 4.

loci (Cronin, 1993; Moritz, 1994b). This has been best demonstrated in certain macaque species (Melnick and Hoelzer, 1993; Hoelzer and Melnick, 1994; see Palumbi and Baker, 1994, for a nonprimate example). Although female transfer in chimpanzees should, in all likelihood, reduce the possibility of observing a pattern of mitochondrial DNA para-

TABLE 4. Frequencies of observed PABX and PABY haplotypes[1]

P. paniscus, observed haplotypes	Number of chromosomes with observed haplotype	Frequency
PABX		
Paniscus A	N = 10	91.0%
Paniscus B	N = 1	9.0%
PABY		
Paniscus A	N = 8	72.7%
Paniscus B	N = 3	27.3%

	Number of chromosomes with observed haplotype[2]			
P. troglodytes, observed haplotypes	*verus*	*troglodytes*	*schweinfurthii*	Frequency
PABX				
Trog 1			1	5%
Trog 2	1		1	10%
Trog 3	2			10%
Trog 4	3			15%
Trog 5	7			35%
Trog 6	4	1		25%
PABY				
Trog 1	16			84%
Trog 2			2	11%
Trog 3		1		5%

[1] The evolutionary relationships among PABX and among PABY haplotypes are presented in Figure 4.
[2] The chimpanzee Bullet was not sequenced for PSUY.

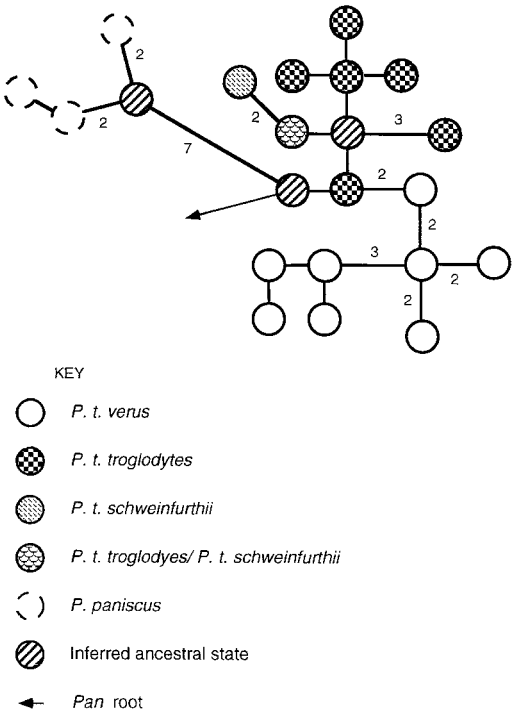

Fig. 5. A maximum parsimony network for all the nuclear data analyzed (i.e., a total (nuclear) evidence approach). Each circle represents an individual(s) (i.e., not haplotypes) and is pattern-coded to indicate the species or subspecies or if they are inferred states. Mutational distances between individuals greater than one are indicated. The *Pan* root is indicated by the arrow and represents the branch leading to the outgroup (i.e., *Gorilla* or *Homo*).

phyly similar to what has been observed in macaques, it is nevertheless important to note that the exact pattern that might be observed remains unknown. Until wild populations of chimpanzees can be adequately sampled for both nuclear and mitochondrial DNA variation, which to date remains difficult (Gagneux et al., 1997), it may be premature to assume that mitochondrial DNA haplotypes "get mixed as extensively as nuclear genes" (Ruvolo, 1994).

A second reason why mtDNA may not be representative of the true genetic relations among chimpanzee populations is that because mitochondria are maternally inherited and are effectively haploid, they have ¼ of the effective population size (N_e) of an autosomal gene. As a result, random drift naturally affects the mitochondrion more than a nuclear locus (Birky et al., 1989). In other words, unique mitochondrial DNA haplotypes have a greater chance of becoming fixed (or lost) in a population than does any autosomal allele. Interestingly, the same reasoning may be used to argue that mitochondrial DNA is a better locus from which to infer evolutionary relationships: mitochondrial DNAs track evolutionary relationships better because the reduced effective population size allow informative characters to become fixed faster (Moore, 1995). The identification of mitochondrial D-loop characters that can be used to subspecifically type

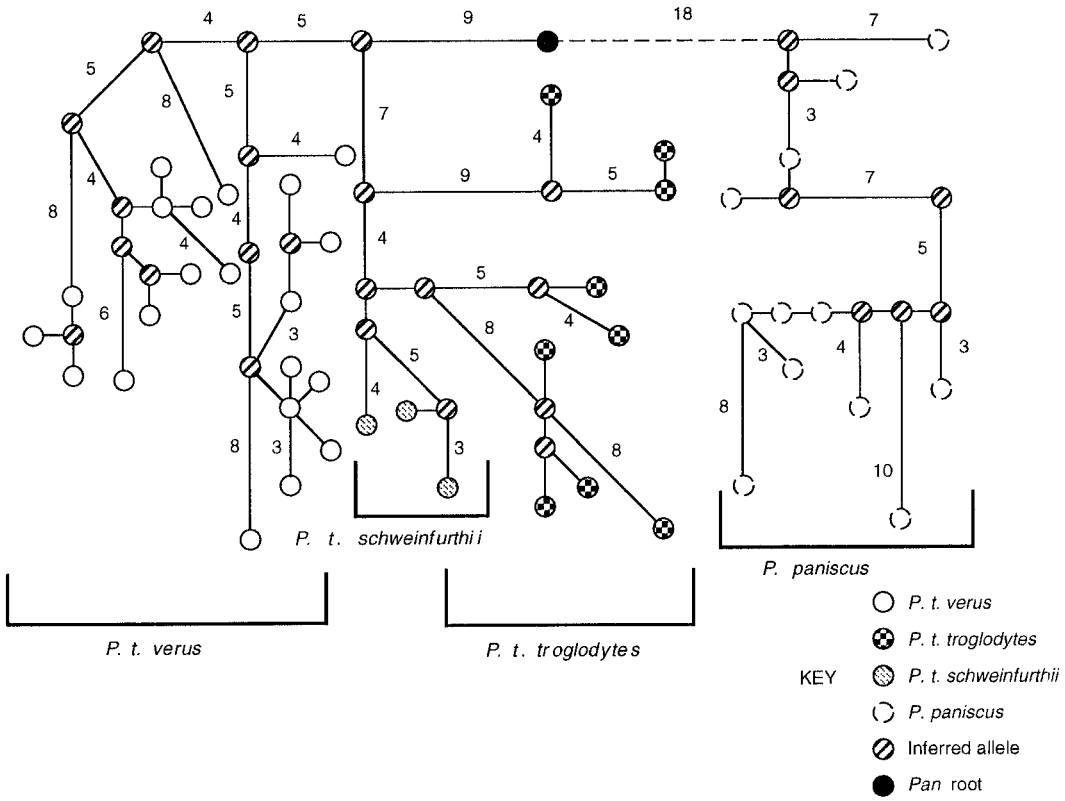

Fig. 6. The maximum parsimony network (representing 1 of at least 6,515 equally parsimonious trees) for the observed mitochondrial D-loop haplotypes. Each circle represents an allele and is pattern-coded to indicate the species or subspecies in which the allele was detected or if it is an inferred ancestral alleles. Mutational distances greater than two are indicated; the branch connecting *P. paniscus* to the *Pan* root is not to scale. It is worth noting that all individuals unambiguously cluster with their known subspecies; this observation is true even if these data are pooled with the data set of Morin et al. (1994) (see Deinard, 1997). The DNA sequences for each observed haplotype have been deposited in Genbank (Accession #AF176709-AF176766).

individual chimpanzees are examples of such "fixed" informative characters (Morin et al., 1992).

Unfortunately, random drift may also "fix" characters that are, in effect, misinformative. This phenomenon may be especially true if comparisons are made between populations that have shared a highly polymorphic common ancestor and have only recently (in an evolutionary sense) become isolated (Lanyon, 1988; Pamilo and Nei, 1988; Moore, 1995), as most certainly the two species and three subspecies of chimpanzees have been. Given the high level of mitochondrial and nuclear polymorphism observed within extant members of the genera (Tables 1–4, Figs. 2–6), it would be difficult to argue that genetic polymorphism was not extensive among the common ancestors of *Pan*. Obviously, the effects of random drift and ancestral polymorphism may also confound the relationships among nuclear alleles, but due to the mitochondrion's faster rate of evolution (Brown et al., 1982) and its maternal inheritance, it becomes easier to imagine chimpanzee mitochondrial DNA lineages sorting such that the observed gene tree disagrees with the subspecies tree (Pamilo and Nei, 1988).

Acknowledging such a possibility, even though it remains to date only a hypothetical concern, is important for chimpanzee conservation and management as it allows for the probability of misclassifying (subspecifically) chimpanzees of unknown geographical origin. This is an important realization, given that any study of wild chimpanzees is automatically provided the

subspecies classifications of the individuals being examined, as chimpanzee subspecies have historically been geographically defined (Hill, 1969). As a result, such studies have the advantage of "knowing" an individual's subspecies classification, irrespective of any discrepancies among the mitochondrial data. The mitochondrial analysis of an individual with an unknown geographical origin (including even some of the individuals sampled for this project), however, would not have this privilege. Although the exact degree of random drift within the chimpanzee mitochondrial genome may not influence our ability to subspecifically classify chimpanzees, it is nevertheless prudent to at least acknowledge the role that such drift could have on classification. This point may be particularly relevant as researchers begin to sample and type the remaining small and isolated populations of chimpanzees (e.g., Gonder et al., 1997).

Lastly, a third reason why mitochondrial DNA (mtDNA) may not be representative of the true genetic relations among populations is that homoplasy affects the mitochondrion more than an autosomal locus. Not only does mitochondrial DNA evolve faster, but because it consists of highly conserved (coding) sequences, it demonstrates both a higher observed substitution rate at third-codon positions (Brown, 1985) and, in general, a large transition bias (Wakeley, 1996). Mitochondrial DNA may therefore experience multiple substitutions at any given nucleotide with a greater frequency than does nuclear DNA. The effect of multiple substitutions may prove to be especially problematic for analyses that limit their comparisons to the rapidly evolving mitochondrial D-loop. Based on the data collected for this project, it is possible to calculate and compare the degree of homoplasy, or similarity that cannot be directly attributed to common ancestry, of the evolutionary networks (Figs. 2–4 and 6) for each locus examined (Table 5), as well as the evolutionary networks that consist of pooled data (all nuclear loci, Fig. 5; all loci, not shown). From these data (Table 5), not only is it obvious that the mitochondrial D-loop data have a much higher level of homoplasy (49%) as compared to the nuclear data (0–11%), but even when all the nuclear data (i.e., from all nuclear loci examined excluding the STR) are combined, the degree of homoplasy remains low (9%). Studies comparing the mitochondrial D-loop of wild chimpanzees have yielded similar results. For example, the data of Morin et al. (1994) contain 46% homoplasy, while the data of Goldberg and Ruvolo (1997) contain 41% homoplasy. Given that homoplasy acts to confound the evolutionary relationships among haplotypes (and hence individuals or populations), "characters" based strictly on mitochondrial data, especially D-loop sequence data, may have a greater chance of being obscured and uninformative than those that are nuclear-based. Consequently, such mitochondrial "characters" may not represent the individual or populational relationships among chimpanzees as accurately as those defined by the nuclear genome.

TABLE 5. Degree of homoplasy within the evolutionary networks examined[1]

Locus	Degree of homoplasy (%)	Σ_{min}	Σ_{tree}
HOXB6	4	26	27
APOB	11	42	47
PABX	0	23	23
PABY	0	20	20
mt *D*-loop	49	200	394
All nuclear	9	72	79
All loci	44	277	499

[1] The degree of observed homoplasy is presented as a percentage. Σ_{min} represents the number of mutational steps that exist for the theoretical shortest tree, while Σ_{tree} represents the number of steps for the shortest obtainable tree. In both cases, the number of steps was calculated from the evolutionary networks (Figs. 2–6) that were rooted by human and/or gorilla sequences that served as outgroup(s). The degree of homoplasy for the "locus" *All nuclear* corresponds to the data set that includes all the nuclear data combined (Fig. 5), while the degree of homoplasy for the "locus" *All loci* corresponds to the data set that includes all data (i.e., nuclear and mitochondrial; not shown).

Another strategy commonly used to generate genetic data on captive (and wild) primates has been the analysis of short tandem repeats (STR or microsatellites). STR loci usually consist of tandem arrays of short (1–5 bp) repeats, and many are considered good markers for generating useful biological data because of their high heterozygosity, their Mendelian inheritance, and in general, the ease with which most can be typed. Recently, researchers have employed STR to infer evolutionary relationships across primate populations (Edwards et al., 1992; Shriver et al., 1995; Goldstein et al., 1995; Deka et al., 1995; Slatkin, 1995; Wise et al.,

1997; Jorde et al., 1997), as well as across primate species (Blanquer-Maumont and Crouau-Roy, 1995; Meyer et al., 1995; Ely et al., 1992; Deka et al., 1994; Kayser et al., 1995; Rubinsztein et al., 1995; Garza et al., 1995).

Unfortunately, there are at least two significant problems with attempting to resolve evolutionary relationships, at either the species or population level, through STR analysis. The first problem concerns the presumption that STR data, when summarized as a distance or a statistic (e.g., Goldstein et al., 1995; Shriver et al., 1995; Slatkin, 1995; Takezaki and Nei, 1996), are appropriate for phylogenetic or genealogical analyses. Such summaries result in the reduction and/or loss of character information because they ignore any hierarchical descent relationships that may be detectable through synapomorphic characters. Thus, this methodology may only provide a relative measure of interspecific relationships. A similar criticism has been leveled at approaches using immunological and electromorph data for phylogeny reconstruction (Farris, 1981).

The second major problem associated with attempts to resolve evolutionary relationships through STR analysis is that the majority of microsatellite studies fail to determine unambiguously whether or not alleles observed to be identical in size are in fact identical by descent, or even identical in sequence. Thus, a high degree of homoplasy may exist within data generated from STR. The HOXB6 STR data illustrate this point. Even though two randomly sampled individuals may share identical STR alleles in repeat length (e.g., 138 bp), they may not share identical nucleotide sequence in the flanking DNA (e.g., Trog 1 and Trog 4; Table 1b). Consequently, the HOXB6 STR can neither provide an accurate representation of the evolutionary relationships among observed alleles (or sampled individuals), nor can it be used to accurately estimate the level of genetic variation between individuals or across populations. Similar conclusions have been made for other STR loci by Garza et al. (1995) and Grimaldi and Crouau-Roy (1997). Although the polymorphic nature of microsatellites allows them to be extremely useful in identifying relationships among closely related individuals (e.g., Jarne and Lagoda, 1996), such as paternity assessments that cannot be determined by mitochondrial analysis, the HOXB6 data should serve as a cautionary example and remind researchers of the uncertainties inherent when attempting to infer evolutionary relationships from STR. The resulting conclusion, therefore, should not be that microsatellite loci are worthless for discriminating evolutionary relationships, but rather that their usefulness may be limited by the amount of information known for any given locus.

CONCLUSIONS

Although the analyses of mitochondrial (D-loop) sequences and microsatellite loci may provide highly polymorphic data that may be used to infer evolutionary relationships and define conservation units among captive chimpanzees, these loci have the potential for a higher degree of homoplasy than do nonrepetitive nuclear loci. Consequently, one may conclude that the analysis of single-copy, noncoding, nonrepetitive nuclear loci may identify conservation units more readily and with less "molecular noise" (e.g., the effects of homoplasy or genetic drift) than comparisons that are limited to either microsatellite or mitochondrial D-loop sequences. It should be emphasized that the nuclear data from this project were generated from three independently evolving nuclear DNA loci, each of which may provide independent data on the evolutionary relationships among the chimpanzees sampled. The pattern of evolutionary relationships supported by these nuclear data can be confirmed because there are thousands of independently evolving nuclear DNA loci. The mitochondrion, on the other hand, represents but a single locus and it is therefore impossible to independently replicate the mitochondrial data.

This conclusion does not suggest that any single nuclear locus is more informative than the mitochondrion, nor does it suggest that any genetic comparison(s) should be limited to either the nuclear or mitochondrial genome. Rather, any comparison will be stronger if multiple independent loci can be compared because each independent locus has its own evolutionary history that may be used to infer the evolutionary relationships among individuals or populations. Given the large number of independent

nuclear loci available for any specific genetic analysis and the fact that there is no independently evolving mitochondrial DNA sequence suggests that the analysis of single-copy, unlinked, noncoding, nonrepetitive nuclear loci for nucleotide variation may provide a greater resource from which to identify conservation management units among captive chimpanzees. Moreover, the analysis of such nuclear loci may become more useful (and more utilized) as data generated from nuclear loci accumulate for both captive and wild chimpanzees. Not only may additional data from well-characterized nuclear loci allow for more precise paternity assessment, but more importantly, such data may allow for the identification of subspecific hybrids within the captive common chimpanzee population. In comparison, rapidly evolving STR loci and the maternally inherited mitochondrion may be unable to contribute to the identification of such hybrid individuals.

Although we note that only a few loci were examined for this project, several additional conclusions may be made concerning these data. First of all, the data suggest that bonobos may carry less genetic variation than do common chimpanzees. Not only do bonobos consistently demonstrate fewer haplotypes at each locus, but they also show marked reductions in nucleotide diversity at the APOB and PABX loci (a 10-fold reduction) and the mitochondrial D-loop locus (a twofold reduction) (Table 2). Given that the results from both Tajima's *D* and Fu and Li's D and F tests argue for mutational neutrality at these loci (Table 2), these results may be explained by either a biased sample of bonobos or a populational bottleneck in the evolutionary past of *P. paniscus*.

Although only additional sampling (either different individuals and/or different loci) will be able to confirm if *P. paniscus* has a genome-wide reduction in nucleotide diversity, it is worth noting that a similar pattern of reduced variation among bonobos has also been observed among chimpanzee blood groups. For example, 16 various V-A-B-D blood types have been observed in common chimpanzees, while only one type (v.D) has been observed in bonobos (Moor-Jankowski and Wiener, 1965; Moor-Jankowski et al., 1975; Socha, 1984). Interestingly, v.D is the rarest V-A-B-D type in common chimpanzees, with a frequency of 1.4% (Socha, 1984). A similar pattern has been observed for the R-C-E-F blood group system: common chimpanzees demonstrate at least 20 different types, while bonobos are monomorphic for an R-C-E-F type not observed in common chimpanzees (Moor-Jankowski et al., 1975; Socha, 1984). From an evolutionary perspective, our data and the blood group data (assuming that blood groups have not experienced positive selection) support the hypothesis that a subpopulation of chimpanzees was isolated (i.e., a bottleneck event) south of the Zaire river and became the "founders" of *P. paniscus*. Whatever the exact cause may be for the pattern of variation observed among bonobos, from a genetic diversity perspective, the suggestion that bonobos may have a reduced degree of genetic diversity should be a concern for those involved with the captive management of *P. paniscus*.

A third conclusion that may be generated from our data is that subspecies clusters can be identified among common chimpanzees. These findings naturally lead to questions surrounding the current management of captive chimpanzees, specifically whether or not common chimpanzee subspecies should be managed independently. For example, the data presented demonstrate that the three *P. t. schweinfurthii* individuals sampled for this project cluster extremely close to, and in some cases with, *P. t. troglodytes*. The lone exception is at PABX, where a *P. t. verus* haplotype and a *P. t. schweinfurthii* haplotype cluster at the root of the *Pan* network (Fig. 4). *P. t. verus*, on the other hand, most often form unique clusters that are mutationally distant from the other two subspecies. An identical pattern was observed from a much larger analysis of mitochondrial DNA (Morin et al., 1994).

What is the significance of this observation? Morin et al. (1994) suggested that the large mitochondrial differentiation between subspecies might be grounds for elevating *P. t. verus* to full species rank (*Pan verus*) if similar results were observed at "three or more unlinked nuclear loci, and supported by eco-behavioral data." We have presented data generated from three unlinked nuclear loci that we believe provide evidence that is more consistent with a subspecific ranking of *P. t. verus*. Specifically, *P. t. verus* is not

distinguished from the other common chimpanzee subspecies as readily as common chimpanzees (*P. troglodytes* ssp.) are distinguished from bonobos (*P. paniscus*). Moreover, the relationships among *Pan* haplotypes (Figs. 2–6) argue that *P. t. verus* are no more or less divergent than any other subspecies with respect to the ancestral *Pan* haplotype (i.e., the *Pan* root). Data from the HOXB6, APOB, and PABX loci illustrate this point: *P. t. troglodytes* represents the most ancestral haplotype at HOXB6, *P. t. verus* represents the most ancestral haplotype at APOB, and both *verus* and *schweinfurthii* represent the most ancestral haplotype at PABX (Tables 1a, 3, 4; Figs. 2–4). Although the lack of differentiation observed in the nuclear data does not specifically exclude the taxonomic shift suggested by Morin et al. (1994), the nuclear data nevertheless do not corroborate the differentiation observed within the mtDNA, that to date remains the only data supporting *Pan verus*.

Consequently, we believe that both the mitochondrial and the nuclear data generated to date support the separate conservation of *P. t. verus*, independent of *P. t. troglodytes* and *P. t. schweinfurthii*, but that these data are consistent with the subspecific classification of *P. t. verus*. Additional nuclear data, from a larger sample of individuals, will need to be collected on *P. t. schweinfurthii* to determine whether or not *schweinfurthii* can be better differentiated from *P. t. troglodytes*. Once this is accomplished, it will be possible to determine whether or not *troglodytes* and *schweinfurthii* represent one or two separate management units (Moritz, 1995). Such questions need to answered sooner rather than later, while "wild-caught" individuals are still in the breeding population. It is only through the further genetic analyses of these "wild-caught" individuals that we will be able to verify the patterns that we are beginning to see from the nuclear genome.

The last conclusion that may be drawn from these data concerns chimpanzee evolution. Because the Miocene fossil record is so poor, there currently exists very little information on the ancestors of extant chimpanzees. The data from this project, however, might provide a rare genetic insight into the origin and evolution of the genus *Pan*. Specifically, we note that for the two loci for which we have generated the most data, the ancestral *Pan* haplotype (i.e., the *Pan* root) is represented by *P. t. troglodytes* (at HOXB6) and *P. t. verus* (at APOB). Although the *Pan* root is represented by a shared *verus*/*schweinfurthii* haplotype at PABX and by an inferred ancestral haplotype at PABY, *P. t. schweinfurthii* do not independently represent the ancestral haplotype at any of the nuclear loci examined for this project, as do *verus* and *troglodytes*. Although we acknowledge that our *schweinfurthii* sample is small, we nevertheless are intrigued by the evidence, albeit limited, that suggest that the geographical origin of *Pan* may be either Central or West Africa, and not East Africa. Although such a hypothesis has previously been suggested (e.g., Kortlandt, 1972), there have been few genetic data from which such a hypothesis could be tested. Obviously, only additional comparisons between chimpanzee subspecies will be able to determine the validity of this claim. However, it is worth noting that the lack of any chimpanzee-like fossils from East Africa would be expected and consistent with *Pan* originating and evolving in Central or West Africa.

ACKNOWLEDGMENTS

We thank Drs. O. Ryder of the San Diego Zoo, E.J. Wickings of the Centre International de Recherches Médicales de Franceville (Franceville, Gabon), A. Prince of the New York Blood Center (New York, NY), D. Povinelli of the New Iberia Research Center (New Iberia, LA), J. Fritz of the Arizona Primate Foundation (Tempe, AZ), and J. Weinberg of Yale University (New Haven, CT), as well as officials of the Milwaukee County Zoo, the Lowery Zoo (Tampa, FL), and the Yerkes Regional Primate Center (Atlanta, GA) for graciously providing blood or DNA samples. Additionally, we thank William Speed for his assistance in DNA sequencing and Rika Kaestle, Andy Clark, and three anonymous reviewers for their helpful suggestions and comments on the manuscript. This research was supported in

APPENDIX A. PCR and DGGE Conditions

Fragment	Primer 1	Primer 2	PCR buffer (mM $MgCl_2$)	Thermal profile (temperatures in °C)	Cycles	Approximate product size
Apolipoprotein B						
A	APOB1cl	APOB3	2.25	30″@94, 30″@56, 30″@72	30	194 bp
B	APOB5	APOB6cl	3.0	30″@94, 30″@59, 30″@72	30	221 bp
C	APOB7	APOB8	3.0	30″@94, 30″@54, 30″@72	30	307 bp
D	APOB9cl	APOB10	2.25	30″@94, 30″@59, 30″@72	30	254 bp
E	APOB11cl	APOB12	3.0	30″@94, 30″@54, 30″@72	30	271 bp
F	APOB4	APOB2.2cl	2.25	30″@94, 30″@58, 30″@72	35	332 bp
Homeobox B6						
A	HOX 88	HOX 292cl	3.0	30″@94, 30″@56, 30″@72	35	209 bp
B	PYG 20cl	PYG 21	2.25	30″@94, 30″@61, 30″@72	35	291 bp
C	PYG 2	PYG 3	2.25	30″@94, 30″@58, 30″@72	30	303 bp
D	PYG22cl	PYG 23	3.75	30″@94, 30″@57.5, 30″@72	35	244 bp
E	Pyg8	Pyg1.mod	2.25	30″@94, 30″@58, 30″@72	30	~136 bp (CA repeat)
F	HOX 860cl	HOX 1146	1.5	30″@94, 30″@58, 30″@72	30	286 bp
Pseudoautosomal boundary						
PSUX.A–PSU.XYR	PSUX.A	PSU.XYR	2.25	30″@94, 30″@58.5, 30″@72	35	769 bp
PSUY.A–PSU.XYR	PSUY.A	PSU.XYR	2.25	30″@94, 30″@58.5, 30″@72	35	945 bp
Mitochondria						
Mitochondrial *D*-loop	D-88	D-441	1.5	30″@94, 30″@61, 30″@72	30	353 bp

	DGGE conditions[1]			
Fragment analyzed	Electrophoretic conditions	DGGE gradient[2]	DGGE gradient[3]	Buffer tank temperature (in °C)
Apolipoprotein B				
A	80 V@20 hr	30–50%	2.1 M/12%–3.5 M/20%	65
F	80 V@20 hr	10–40%	0.7 M/4%–2.8 M/16%	65
B	150 V@15 hr	30–50%	2.1 M/12%–3.5 M/20%	60
C	133 V@17 hr	30–50%	2.1 M/12%–3.5 M/20%	60
D	150 V@15 hr	30–50%	2.1 M/12%–3.5 M/20%	60
E	150 V@18 hr	10–60%	0.7 M/4%–3.2 M/24%	60
Homeobox B6				
A	150 V@15 hr	50–80%	3.5 M/20%–5.6 M/32%	60
B	150 V@15 hr	30–80%	2.1 M/12%–5.6 M/32%	60
C	150 V@15 hr	30–60%	2.1 M/12%–3.2 M/24%	60
D	150 V@15 hr	30–70%	2.1 M/12%–4.9 M/28%	60
F	113 V@17 hr	20–80%	1.4 M/8%–5.6 M/32%	55

[1] DGGE was performed with 8% acrylamide with 1× Tris-acetate-ethylenediaminetetraacetate (TAE) buffer, with the gel assembly immersed in a 1× TAE buffer tank at 55–65°C (constant).
[2] In terms (%) of a reference acrylamide solution of a reference solution of 7 M urea/40% (v/v) formamide.
[3] In terms of urea (in M)/formamide (in %) denaturants.

part by grants from the Leakey Foundation, the Wenner-Gren Foundation for Anthropological Research, and the National Science Foundation (SBR 9315871) to A.S.D.

APPENDIX B: ORIGIN OF SAMPLES

The chimpanzees sampled for this project, and the researchers and/or institutions providing the samples, were as follows: *Pan troglodytes troglodytes*—Cheetah, Dodo, Berthe, Gemeni, Bakoumba, Julie, Noemie, N'tebe, and Cleo (E.J. Wickings, Centre Internationale de Recherches Médicales de Franceville, Franceville, Gabon); *Pan troglodytes verus*—Sara, Mabel, Ettie, Priapus, Lowie, Rinus, Brian, Anita, Hannibal, Lottie, Juno, and Bullet (A. Prince, New York Blood Center and the Vialab Research Facility, Robertsfiled, Liberia), and Vincent, Jiggs, Harry, Gus, Bert, Ted, Tate, Dude, A136, A202, A208, A237, A248, A263, A333, and A336 (New Iberia Research Center, New Iberia, LA) and Herman (Lowery Zoo); *Pan troglodytes schweinfurthii*—Harriet and Kobi (J. Fritz, Arizona Primate Foundation) and Mgbadolite (E.J. Wickings, Centre Internationale de Recherches Médicales de Franceville); *Pan paniscus*—Kakowet, Lady, Linda, Vernon, Joey, Bosondjo, Kitty, Matata, Jimmy, Toby, Susie, Congo, Maiko, Charlie (O. Ryder), Lody, Maringa (Milwaukee Zoo), Kidongo (Yerkes Regional Primate Center),

and Diatou (J. Weinberg). All DNA samples used for this study were extracted from lymphoblastoid cell lines established within our laboratory using a standard phenol-chloroform protocol, except for the DNA samples provided by O. Ryder. DNA was also extracted from whole blood from the following individuals for which cell lines do not exist: Cleo, Berthe, Gemeni, and N'tebe. Although no geographical information (and hence no subspecies classification) was available for the individuals housed at the New Iberia Research Center, they may be regarded, conditionally, as *P. t. verus* based upon their mitochondrial D-loop sequences (Morin et al., 1994); the nuclear sequences generated from these samples do not contradict this classification. Importation of the samples provided by the Centre Internationale de Recherches Médicales de Franceville was carried out under a CITES permit.

LITERATURE CITED

Arnason U, Gullberg A, Janke A, Xu X. 1996. Pattern and timing of evolutionary divergences among hominoids based on analyses of complete mtDNAs. J Mol Evol 43:650–661.

Avise J. 1995. Mitochondrial DNA polymorphism and a connection between genes and demography of relevance to conservation. Conserv Biol 9:686–690.

Birky C, Fuerst P, Maruyama T. 1989. Organelle gene diversity under migration, mutation and drift: equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. Genetics 121:613–627.

Blanquer-Maumont A, Crouau-Roy B. 1995. Polymorphism, monomorphism and sequences in conserved microsatellites in primate species. J Mol Evol 41:492–497.

Brown W. 1985. The mitochondrial genome of animals. In: MacIntyre R, editor. Molecular evolutionary genetics. New York: Plenum Press. p 95–130.

Brown W, Prager E, Wang A, Wilson A. 1982. Mitochondrial DNA sequences of primates: tempo and mode of evolution. J Mol Evol 18:225–239.

Chesser R, Smith M, Brisbin I. 1980. Management and maintenance of genetic variability in endangered species. Int Zoo Ybk 20:146–154.

Cronin M. 1993. Mitochondrial DNA in wildlife taxonomy and conservation biology: cautionary notes. Wildlife Soc Bull 21:339–348.

Deinard A. 1997. The evolutionary genetics of the chimpanzees. Ph.D. thesis, Yale University.

Deinard A, Kidd K. 1995. Levels of DNA polymorphism in extant and extinct hominoids. In: Brenner S, Hanihara K, editors. The origin and past of modern humans as viewed from DNA: proceedings of the workshop on the origin and past of *Homo sapiens sapiens* as viewed from DNA—theoretical approach. Singapore: World Scientific. p 149–170.

Deinard A, Kidd K. 1998. Evolution of a D2 dopamine receptor intron within the great apes and humans. DNA Seq 8:289–301.

Deinard A, Ruano G, Kidd K. 1992. A dinucleotide repeat polymorphism at the HOX2B locus. Nucleic Acids Res 20:1171.

Deka R, Shriver M, Yu L, Jin L, Aston C, Chakraborty R, Ferrell R. 1994. Conservation of human chromosome 13 polymorphic microsatellite $(CA)_n$ repeats in chimpanzees. Genomics 22:226–230.

Deka R, Jin L, Shriver M, Yu L, DeCroo S, Hundrieser J, Bunker C, Ferrell R, Chakraborty R. 1995. Population genetics of dinucleotide $(dC\text{-}dA)_n \cdot (dG\text{-}dT)_n$ polymorphism in world populations. Am J Hum Genet 56:461–474.

Edwards A, Hammond H, Jin L, Caskey T, Chakraborty R. 1992. Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. Genomics 12:241–253.

Ely J, Deka R, Chakraborty R, Ferrell R. 1992. Comparison of five tandem repeat loci between humans and chimpanzees. Genomics 14:692–698.

Farris JS. 1981. Distance data in phylogenetic analysis. In: Funk V, Brooks DR, editors. Advances in cladistics: proceedings of the first meeting of the Willi Hennig Society. New York: New York Botanical Society. p 3–23.

Fischer S, Lerman L. 1983. DNA fragments differing by a single base substitution are separated in denaturing gradient gels: correspondence with melting theory. Proc Natl Acad Sci USA 80:1579–1583.

Fisher R. 1930. The genetic theory of natural selection. Oxford: Oxford University Press.

Fu YX, Li WH. 1993. Statistical tests of neutrality of mutations. Genetics 133:693–709.

Gagneux P, Boesch C, Woodruff D. 1997. Microsatellite scoring errors associated with noninvasive genotyping based on nuclear DNA amplified from shed hair. Mol Ecol 6:861–868.

Garza J, Slatkin M, Freimer N. 1995. Microsatellite allele frequencies in humans and chimpanzees, with implications for constraints on allele size. Mol Biol Evol 12:594–603.

Goldberg T, Ruvolo M. 1997. Molecular phylogenetics and historical biogeography of east African chimpanzees. Biol J Linn Soc 61:301–324.

Goldstein D, Linares A, Cavalli-Sforza L, Feldman W. 1995. Genetic absolute dating based on microsatellites and the origin of modern humans. Proc Natl Acad Sci USA 92:6723–6727.

Gonder MK, Oates J, Disotell T, Forstner M, Morales J, Melnick D. 1997. A new West African chimpanzee subspecies? Nature 388:337.

Grimaldi M, Crouau-Roy B. 1997. Microsatellite allelic homoplasy due to variable flanking sequences. J Mol Evol 44:336–340.

Hill WO. 1969. The nomenclature, taxonomy and distribution of chimpanzees. In: Bourne G, editor. The chimpanzee: anatomy, behavior, and diseases of chimpanzees. Basel: Karger. p 22–49.

Hoelzer G, Melnick D. 1994. Patterns of speciation and limits to phylogenetic resolution. Trends Ecol Evol 9:104–107.

Hoelzer G, Dittus W, Ashley M, Melnick D. 1994. The local distribution of highly divergent mitochondrial DNA haplotypes in toque macaques *Macaca sinica* at Polonnaruwa, Sri Lanka. Mol Ecol 3:451–458.

Horai S, Hayasaka K, Kondo R, Tsugane K, Takahata N. 1995. Recent African origin of modern humans revealed by complete sequences of hominoid mitochondrial DNAs. Proc Natl Acad Sci USA 92:532–536.

Huelsenbeck J, Bull J, Cunningham C. 1996. Combining data in phylogenetic analysis. Trends Ecol Evol 11:152–158.

James R, Lebrg P, Quattro J, Vrijenhoek R. 1997. Genetic diversity in black howler monkeys (*Alouatta pigra*) from Belize. Am J Phys Anthropol 102:329–336.
Jarne P, Lagoda P. 1996. Microsatellites, from molecules to populations and back. Trends Ecol Evol 11:424–429.
Jorde L, Rogers A, Bamshad M, Watkins WS, Krakowiak P, Sung S, Kere J, Harpending H. 1997. Microsatellite diversity and the demographic history of modern humans. Proc Natl Acad Sci USA 94:3100–3103.
Kano T. 1986. The last ape: pygmy chimpanzee behavior and ecology. Translated by E. Ono Vineberg, 1992. Stanford: Stanford University Press.
Kayser M, Nurnberg P, Bercovitch F, Nagy M, Roewer L. 1995. Increased microsatellite variability in *Macaca mulatta* compared to humans due to a large scale deletion/insertion event during primate evolution. Electrophoresis 16:1607–1611.
Kidd K, Ruano G. 1992. PCR techniques for studies of DNA sequence variation. Crime Lab Digest 18.
Kortlandt A. 1972. New perspectives on ape and human evolution. Amsterdam: Stichting Voor Psychobiologie.
Lacy R. 1987. Loss of genetic diversity from managed populations: interacting effects of drift, mutation, immigration, selection, and population subdivision. Conserv Biol 1:143–158.
Lanyon S. 1988. The stochastic mode of molecular evolution: what consequences for systematic investigations? Auk 105:565–573.
Lerman L, Fischer S, Hurley I, Silverstein K, Lumelsky N. 1984. Sequence-determined DNA separations. Annu Rev Biophys Bioeng 13:399–423.
Melnick D, Hoelzer G. 1993. What is mtDNA good for in the study of primate evolution? Evol Anthropol 2:2–10.
Meyer E, Rand S, Kuhlmann D, Brack M, Brickmann B. 1995. Microsatellite polymorphisms reveal phylogenetic relationships in primates. J Mol Evol 41:10–14.
Mitton J. 1997. Selection in natural populations. Oxford: Oxford University Press.
Moore W. 1995. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. Evolution 49:718–726.
Moor-Jankowski J, Wiener A. 1965. Simian-type blood group factors in non-human primates. Nature 205: 369–371.
Moor-Jankowski J, Wiener A, Socha W, Gordon E, Mortlemans J, Sedgwick C. 1975. Blood groups of pygmy chimpanzees (*Pan paniscus*): human-type and simian-type. J Med Primatol 4:262–267.
Morin P, Moore J, Woodruff D. 1992. Identification of chimpanzee subspecies with DNA from hair and allele-specific probes. Proc R Soc Lond [Biol] 249:292–297.
Morin P, Moore J, Chakraborty R, Jin L, Goodall J, Woodruff D. 1994. Kin selection, social structure, gene flow, and the evolution of chimpanzees. Science 265: 1193–1201.
Moritz C. 1994a. Defining "evolutionarily significant units" for conservation. Trends Ecol Evol 9:373–375.
Moritz C. 1994b. Applications of mitochondrial DNA analysis in conservation: a critical review. Mol Ecol 3:401–411.
Moritz C. 1995. Uses of molecular phylogenies for conservation. Philos Trans R Soc Lond [Biol] 349:113–118.
Myers R, Maniatis T. 1986. Recent advances in the development of methods for detecting single-base substitutions associated with human genetic diseases. Cold Spring Harbor Symp Quant Biol 51:275–284.
Myers R, Maniatis T, Lerman L. 1987. Detection and localization of single base changes by denaturing gradient gel electrophoresis. Methods Enzymol 155: 501–527.
Myers R, Sheffield V, Cox D. 1989. Mutation detection by PCR, GC-clamps, and denaturing gradient gel electrophoresis. In: Erlich H, editor. PCR technology: principles and applications for DNA amplification. New York: Stockton Press. p 71–88.
Nei M. 1987. Molecular evolutionary genetics. New York: Columbia University Press.
Palumbi S, Baker C. 1994. EPIC amplifications of nuclear DNA introns: opposing views of population structure using mitochondrial and nuclear sequences from humpback whales. Mol Biol Evol 11:426–435.
Pamilo P, Nei M. 1988. Relationships between gene trees and species trees. Mol Biol Evol 5:568–583.
Reinartz G. 1994. Bonobo (*Pan paniscus*) species survival plan 1994–1996. Milwaukee: Zoological Society of Milwaukee County.
Rozas J, Rozas R. 1997. DNAsp version 2.0: a novel software package for extensive molecular population genetics analysis. Comput Appl Biosci 13:307–311.
Ruano G, Kidd K. 1992. Modeling of heteroduplex formation during PCR from mixtures of DNA templates. PCR Methods Amplifications 2:112–116.
Ruano G, Rogers J, Ferguson-Smith A, Kidd K. 1992. DNA sequence polymorphism within hominoid species exceeds the number of phylogenetically informative characters for a HOX2 locus. Mol Biol Evol 9:575–586.
Rubinsztein D, Leggo J, Amos W. 1995. Microsatellites evolve more rapidly in humans than in chimpanzees. Genomics 30:610–612.
Ruvolo M. 1994. Molecular evolutionary processes and conflicting gene trees: the hominoid case. Am J Phys Anthropol 94:94–113.
Ruvolo M, Pan D, Zehr S, Goldberg T, Disotell T, Von Dornum M. 1994. Gene trees and hominoid phylogeny. Proc Natl Acad Sci USA 91:8900–8904.
Saiki R, Scharf S, Faloona F, Mullis K, Horn G, Erlich H, Arnheim N. 1985. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for the analysis of sickle cell anemia. Science 230:1350–1354.
Sambrook J, Fritsch E, Maniatis T. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.
Schobert L. 1991. North American regional chimpanzee studbook, Asheboro, NC no. 2 supplement. North Carolina Zoological Park.
Schonewald-Cox C, Chambers S, MacBryde B, Thomas W, editors. 1983. Genetics and Conservation: a reference for managing wild animal and plant populations. London: Benjamin/Cummings Publishing.
Shriver M, Jin L, Boerwinkle E, Deka R, Ferrell R, Chakraborty R. 1995. A novel measure of genetic distance for highly polymorphic tandem repeat loci. Mol Biol Evol 12:914–920.
Slatkin M. 1995. A measure of population subdivision based on microsatellite allele frequencies. Genetics 139:457–462.
Socha W. 1984. Blood groups of pygmy and common chimpanzees. In: Susman R, editor. The pygmy chimpanzee. New York: Plenum Press. p 13–41.
Susman R, editor. 1984. The pygmy chimpanzee. New York: Plenum Press.
Swofford D. 1990. PAUP: phylogenetic analysis using parsimony, version 3.0. Computer program distributed by the Illinois Natural History Survey, Champaign.
Tagliaro C, Schneider M, Schneider H, Sampaio I,

Stanhope M. 1997. Marmoset phylogenetics, conservation perspectives, and evolution of the mtDNA control region. Mol Biol Evol 14:674–684.

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585–595.

Takezaki N, Nei M. 1996. Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. Genetics 144:389–399.

Wakeley J. 1996. The excess of transitions among nucleotide substitutions: new methods of estimating transition bias underscore its significance. Trends Ecol Evol 11:158–163.

Wise C, Sraml M, Rubinsztein D, Easteal S. 1997. Comparative nuclear and mitochondrial genome diversity in humans and chimpanzees. Mol Biol Evol 14:707–716.

Zhi L, Karesh W, Janczewski D, Frazier-Taylor H, Sajuthi D, Gombeck F, Andau M, Martenson J, O'Brien S. 1996. Genomic differentiation among natural populations of orang-utan (*Pongo pygmaeus*). Curr Biol 6:1326–1336.